AD NUMBER

AD003247

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution: No Foreign.

AUTHORITY

AFAL ltr., 17 Aug 1979

WADC TECHNICAL REPORT 52-224

RETENTION OF TRANSFER IN MOTOR LEARNING AFTER 24 HOURS AND AFTER 14 MONTHS AS A FUNCTION OF DEGREE OF FIRST-TASK LEARNING AND INTER-TASK SIMILARITY

CARL P. DUNCAN
BENTON J. UNDERWOOD

NORTHWESTERN UNIVERSITY

OCTOBER 1952

WRIGHT AIR DEVELOPMENT CENTER

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The information furnished herewith is made available for study upon the understanding that the Government's proprietary interests in and relating thereto shall not be impaired. It is desired that the Judge Advocate (WCJ), Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, be promptly notified of any apparent conflict between the Government's proprietary interests and those of others.

WADC TECHNICAL REPORT 52-224

RETENTION OF TRANSFER IN MOTOR LEARNING AFTER 24 HOURS AND AFTER 14 MONTHS AS A FUNCTION OF DEGREE OF FIRST-TASK LEARNING AND INTER-TASK SIMILARITY

Carl P. Duncan
Benton J. Underwood

Northwestern University

October 1952

Aero Medical Laboratory
Contract No. 33(038)-11396
RDO No. 694-44

Wright Air Development Center
Air Research and Development Command
United States Air Force
Wright-Patterson Air Force Base, Ohio

McGregor & Werner, Inc.
Dayton, O. 150, 17 Feb. 53

FOREWORD

This report was prepared by Northwestern University under USAF Contract No. AF 33(038)-11396. The contract was initiated under a project identified by Research and Development Order 694-17, "Design and Arrangement of Aircraft Controls," and later continued under Research and Development Order 694-44, "Learning and Transfer in Reference to Training Aid Design." The contract was administered by the Psychology Branch of the Aero Medical Laboratory, Directorate of Research, Wright Air Development Center, with Gordon A. Eckstrand acting as Project Engineer.

ABSTRACT

Retention of a transfer task provided on a self-paced discriminative motor device was studied as a function of degree of learning of the training task and similarity between tasks. Retention of the transfer task was measured after 24 hours and again after 14 months following acquisition.

There was some forgetting over 24 hours but in relearning the positive transfer obtained during acquisition of the transfer task continued to be manifest and to vary directly both with degree of first-task learning and with task similarity. Proactive facilitation of retention was obtained.

Forgetting over 14 months was great and showed evidence for differential proactive inhibition as a function of degree of learning. Relearning proceeded relatively rapidly. Performance during relearning varied directly with degree of first-task learning but did not vary with inter-task similarity.

PUBLICATION REVIEW

This report has been reviewed and is approved.

FOR THE COMMANDING GENERAL:

Robert H. Blount

ROBERT H. BLOUNT
Colonel, USAF (MC)
Chief, Aero Medical Laboratory
Directorate of Research

Table of Contents

RETENTION OF TRANSFER IN MOTOR LEARNING AFTER 24 HOURS AND AFTER 14 MONTHS AS A FUNCTION OF DEGREE OF FIRST-TASK LEARNING AND INTER-TASK SIMILARITY

I. INTRODUCTION

In an experiment previously reported (3) the effect of two variables on transfer in motor learning was studied: degree of learning of the first task and similarity between the first and the second (transfer) task. In the present study data will be reported for retention and relearning of the second task measured both 24 hours and 14 months (average) after original learning of the second task.

As far as is known, there are no other studies that have any direct bearing on the present one. It is true that measuring retention and relearning of the second of two successively-learned tasks is the standard design for studying proactive inhibition of recall (7). However, in most such studies proactive inhibition is expected to occur only when the two tasks are mutually interfering; i.e., there is negative transfer from the first to second tasks (7). It has already been shown in the first report (3) that the two motor tasks used in the present research transfer positively. Not only was there no evidence of negative transfer but data reported at that time for retention of the second task after 24 hours showed proactive facilitation, not proactive inhibition, from the first task.

Studies of retention over long periods of time are reports of the relearning of single tasks (5) and so are not relevant to the present research. In the area of transfer, Bunch (see, for example, 1 and 2) has varied the time interval between original learning of the training and transfer tasks in several studies, thereby studying the retention of transfer from the first task measured during acquisition of the second task. The present study differs in that the retention of transfer is studied by measuring recall and relearning of the second task.

II. PROCEDURE

Since this is a study of retention by a portion of the same subjects of the same task described in the first report, details of procedure

concerning subjects, variables, tasks, and apparatus can be obtained from that report (3). Briefly, the apparatus required that the subject learn to move a lever, held by the right hand, into six radially-arranged slots in response to six colored-light stimuli. Simultaneously, the subject held another lever steady with the left hand. If the left lever were not held in the correct position, correct responses (movements of the right lever into the slot which was paired with the stimulus light showing) could not be recorded. Thus, essentially the apparatus provided a paired-associates task with the right lever and a steadiness task with the left lever. Each trial was 20 seconds with 10 seconds inter-trial rest. Separate counters recorded movements of the right lever as correct responses or errors and a clock recorded time that the left lever was held in correct position. With this apparatus the first task for each subject is a particular pairing of the stimulus lights and slots; the second task is provided by differently pairing the lights and slots. The particular light-slot combinations chosen for each subject for his first and second tasks were such that, as explained more fully in the first report, the two tasks are assumed to be not only equal in difficulty but in a sense identical. This has the advantage that data from the first task can be used as control for the second task; separate control groups practicing only the second task were not run.

The first report was concerned with transfer to a second task as a function of two variables: degree of first-task learning and inter-task similarity. There were four degrees of the learning variable: 10, 40, 80, and 180 trials on the first task. The Learning Groups will be denoted by the Roman numerals I, II, III, and IV, with Group I being the 10-trial group. Similarity between tasks was defined in terms of the number of light-slot connections that were different on the second task from what they had been on the first. The three degrees of similarity were obtained by newly pairing 2, 4, or all 6 lights and slots for the second task. The Similarity Groups will be denoted by the capital letters A, B, and C, with Group A being the high similarity group (two lights newly paired on the second task, the other four remaining the same). These two variables were studied in a 4 x 3 factorial design with 300 subjects, 25 in each of the 12 cells.

Two minutes after completion of the first task each subject was given 60 trials on the second task. Twenty-four hours later all 300 subjects were given 20 relearning trials on the second task.

The first report presented the data for acquisition of the second task and for relearning of the same task after 24 hours by all 300 subjects. The present report will present data for relearning of the second task after 24 hours and again after an average of 14 months for only those 186 subjects out of the 300 who could be obtained for retesting the following year.

The 186 subjects on whom the present report is based were run originally between November 7, 1950 and May 18, 1951. As explained above each subject was trained on the first task, trained on the second

task, and tested for recall and relearning of the second task after a 24-hour rest, all of which took from two to five days depending upon the amount of first-task training. All 186 were given the long-term retention test between April 8 and June 7, 1952, with all but three subjects being run in April and May. Table 1 shows the mean and the range in days elapsed between the 24-hr. retention test and the 14-month test for each of the 12 subgroups.

Table 1 - The Mean and the Range of Days Elapsed Between the 24-Hour Retention Test and the Long-term Retention Test of the Transfer Task. Roman Numerals Indicate Degrees of First-Task Learning; Capital Letters Indicate Degrees of Inter-task Similarity.

Mean Days

Group	A	B	C	Grand Mean
I	420.5	429.9	421.3	423.9
II	415.8	427.7	397.6	413.7
III	436.9	422.4	423.1	427.5
IV	410.3	430.3	403.1	414.6
Grand Mean	420.9	427.6	411.3	419.9

Range of Days

Group	A	B	C	Total Range
I	357-537	362-531	341-530	341-537
II	362-530	348-519	364-444	348-530
III	369-562	364-543	343-537	343-562
IV	350-511	349-543	348-530	348-543
Total Range	350-562	348-543	341-537	341-562

The 300 subjects originally used were male undergraduates at Northwestern University. Most of the 186 who returned for the long-term retention test were still undergraduates on the campus in 1952. A few had graduated but were still living in the Chicago area. The number of subjects in each subgroup who returned to be tested for relearning in 1952 is shown in Table 2. Table 2 shows, therefore, the number of subjects on whom the present report is based.

Table 2 - The Number of Subjects in Each Group

Group	A	B	C	Total
I	15	17	16	48
II	17	15	15	47
III	15	16	15	46
IV	15	15	15	45
Total	62	63	61	186

When a subject came in for the relearning session he was told that he was going to have exactly the same task, the same light-slot relationship, that he had been working on when he finished the experiment the previous year. He was then given 20 trials and, as before, during the inter-trial rests he was told how many correct responses he had made in the previous trial.

III. RESULTS

Acquisition of the Second Task

Correct Responses: The major results are presented in Figs. 1 and 2 where correct responses per trial are plotted as a function of pairs of trials. In Fig. 1 the three Similarity Groups have been combined within each of the four Learning Groups to show the effect of first-task learning with similarity held constant. Fig. 2 shows the three

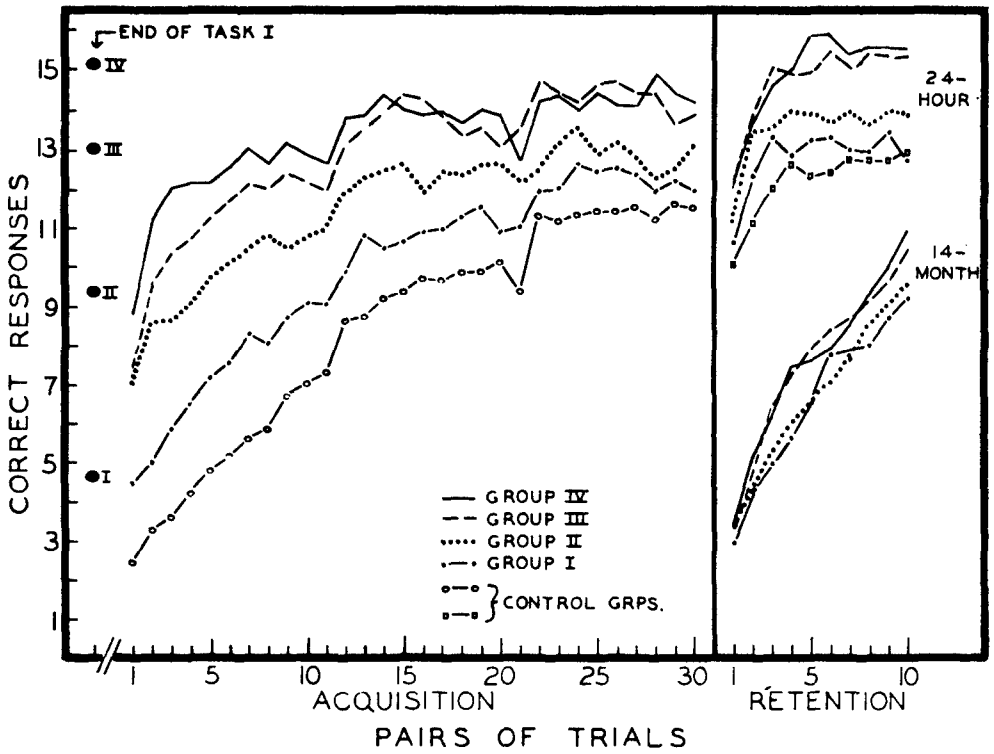

Fig. 1: Acquisition and retention of the transfer task for the four groups having different degrees of learning of the first task. Inter-task similarity is held constant. See text for explanation of this and the following figure.

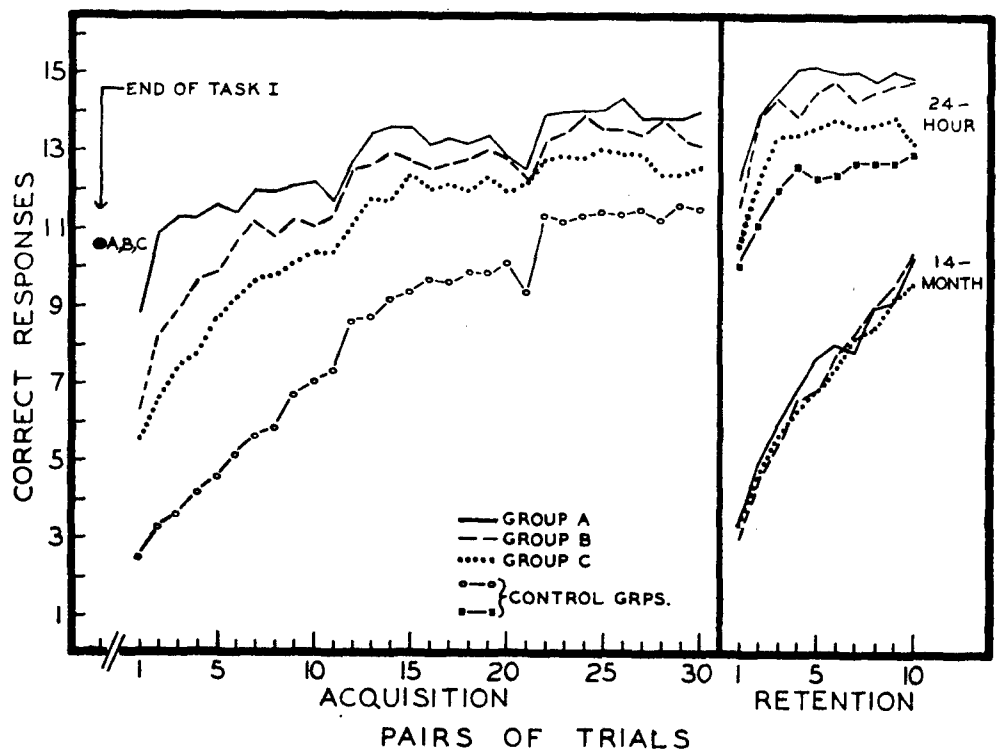

Fig. 2: Acquisition and retention of the transfer task for the three groups having different degrees of inter-task similarity. Degree of first-task learning is held constant.

Similarity Groups with learning constant. Reading from left to right these graphs show:

1. Heavy points indicating performance at the end of training on the first task (mean of the last two trials). Since similarity was not a variable on the first task, Fig. 2 shows a single point, the grand mean for all subjects.

2. Acquisition of the second or transfer task. In both Figs. 1 and 2 the lowest curve in this portion of the graphs is the control for acquisition of the second task. As explained above, and more fully in the first report (3), this control curve is plotted from some of the Task I data. Specifically, the curve shows performance on the first 60 trials on the first task by Groups III and IV combined (N=91). These were the only two groups given at least 60 trials on Task I.

3. At the top of the figures, relearning of the second task after 24 hours, and at the bottom, relearning after an average of 14 months. The lowest curve among the group of curves for 24-hour retention is a control. Again, this control curve consists of data from the first task. Specifically, the curve is plotted from the scores of Group IV (N=45) on trials 61-80 of the first task. This group had their first 60 trials on the first task (the same number of trials given all subjects on the second task) on one day, followed by a 24-hour rest before continuing with trial 61. Therefore their trials 61-80 on the following day show, in effect, relearning of the second task after 24 hours by a group that had no practice on a first task.

We may first point out that the 186 subjects who were available for the long-term retention test are closely representative of the original 300 subjects. On performance at the end of the first task, on acquisition of the second task, and on retention of the second task after 24 hours, the curves of Figs. 1 and 2 are highly similar in both slope and absolute level to the results for all 300 subjects, as comparison of the corresponding figures (Figs. 9 and 10, 14 and 15) in the first report (3) will show. Furthermore, the analyses to be presented below for acquisition and 24-hour retention of the second task by the 186 subjects of the present study will be seen to result in much the same conclusions as were drawn for all 300 subjects.

Inspection of Figs. 1 and 2 shows that there was positive transfer from the first task to acquisition of the second task. No test of significance appears necessary; all experimental curves are above the control curve (the curve for acquisition of the first task). It would also appear that there is differential positive transfer as a function of both learning and similarity since the experimental group curves are clearly different from each other in both Figs. 1 and 2, at least in in the early part of acquisition. We may, however, test the two variables for significance and simultaneously determine if the variables interact. Table 3 summarizes the results of an analysis of variance

on the total correct responses for all 60 acquisition trials on the second task. (It was first determined that the variances of the 12 subgroups were not heterogeneous; chi-square by Bartlett's test was 11.68, where 19.68 is needed for significance at P = .05 for 11 df).

Table 3 - Analysis of Variance on the Total Correct Responses for all 60 Acquisition Trials on the Second Task

Source	df	Mean Square	F
Learning	3	404,945.27	22.08
Similarity	2	179,375.50	9.78
L x S	6	819.12	
Within Groups	174	18,338.61	

F = 2.67 at the 5% level, 3.91 at the 1% level, for 3 and 150 df.

F = 3.06 at the 5% level, 4.75 at the 1% level, for 2 and 150 df.

Table 3 shows that there was no significant interaction between the two variables. Using the variance within groups as error, both learning and similarity are highly significant; it can be concluded that differential positive transfer was produced by the different degrees of each variable.

The analysis presented in Table 3, utilizing, as it does, all 60 trials, tends to mask possible changes in effectiveness of the two variables over practice. As shown in Figs. 1 and 2, during the last 30 acquisition trials the curves for the four Learning Groups are still at somewhat different levels but the curves for the three Similarity Groups overlap considerably. To test possible changes over practice, two analyses of variance were performed, one on the total correct responses for the first 30 trials, and one on the total for the last 30 trials. In both cases it was assumed, without testing, that the variances of the 12 subgroups were sufficiently homogeneous to permit analysis. Tables for the analyses will not be presented; for both the df's and the values required for significance are the same as those given in Table 3.

Analysis of the first 30 trials showed learning highly significant (F = 36.88), similarity highly significant (F = 18.68) and interaction not significant (F = 1.22, where 2.16 is required for the 5% level for 6 and 150 df). Analysis of the last 30 trials showed learning highly significant (F = 10.46), similarity barely significant (F = 3.07), and interaction not significant. The finding that the Similarity Groups were significantly different for the last 30 trials, although at the 5% level only, is a disagreement with the results for the original 300 subjects. For this total group, similarity was not significant during the last 30 acquisition trials on the second task. Nevertheless, the conclusion that was drawn for the original 300 subjects seems at least partly true for the present 186 subjects: the effect of varying first-task learning lasts later during second-task practice than does the effect of varying inter-task similarity.

Errors and Time Scores: There were two types of errors: partial entries of the right-hand lever into an incorrect slot and complete entries into an incorrect slot. These deep errors were complicated by the fact that a complete entry into a correct slot at a time when the left-hand lever was not being held in the correct position also resulted in deep errors being recorded. Because it is a conglomerate, the deep error score will not be analyzed. We shall, however, make use of the shallow error entries into incorrect slots, which from here on will simply be called errors.

The time score represents the time in seconds during which the left-hand lever was held in the correct position during each 20-second trial; the maximum score is therefore 20 seconds. Since the subject could not record correct responses with the right lever whenever the left lever was out of position, the time data will be presented whenever necessary for comparison with the correct responses curves.

In this section on acquisition of the second task there is no need to present time or error data. Differences still existing among the various groups at the end of the acquisition period will be shown in the appropriate figures in the sections below on retention.

Retention of the Second Task after 24 Hours

Correct Responses: Recall and relearning of the second task is depicted by the curves in the upper right-hand portion of Figs. 1 and 2. It can be seen that there was forgetting in all groups. A measure of forgetting was provided by subtracting the score of the first relearning trial from the mean score on the last ten acquisition trials. These losses are shown in Table 4.

Table 4 - Forgetting of the Second Task over a 24-Hour Rest Measured as the Difference Between the Mean Correct Responses of the Last Ten Trials before Rest and the First Trial after Rest.

Learning Group	Mean score before rest	Mean score after rest	Difference
I	12.09	9.92	2.17
II	12.79	10.32	2.47
III	14.18	10.90	3.28
IV	14.44	11.66	2.78
Similarity Group			
A	13.88	11.53	2.35
B	13.53	10.71	2.82
C	12.64	9.80	2.84
Control Group	11.14	9.31	1.83

From Table 4 it can be seen that except for Group III, which is out of line, the amount forgotten increased directly as first-task learning increased and as similarity decreased. It may also be seen that all experimental groups forgot more than the control, i.e., the values suggest proactive inhibition of retention. The forgetting data will first be analyzed for significance of differences among the experimental groups and then comparisons will be made with the control group to determine if there is significant proactive inhibition.

The test for differential forgetting among the experimental groups was by analysis of variance, the results of which are shown in Table 5. It may be seen that the interaction term is significant, probably due to Group III. No particular importance will be attached to the significant interaction. In the previous report (3) the same analysis, based on 300 subjects did not yield a significant interaction, and there was no reversal among forgetting scores as appears in the present Table 4. Table 5 also shows that neither learning nor similarity is significant, regardless of whether interaction or within groups is used as error (F for learning, using within groups as error, is 1.38). We conclude there was no differential forgetting among experimental groups.

Table 5 - Analysis of Variance
on Forgetting over the 24-hour Rest

Source	df	Mean Square	F	P
Learning	3	10.58		
Similarity	2	4.62		
L x S	6	28.58	3.73	<.01
Within Groups	174	7.66		

Since there were no differences among experimental groups in forgetting over 24 hours, we could combine all experimental groups and make a single comparison with the control group to test for proactive inhibition. However, the means compared would be based on greatly different N's, 186 experimental subjects versus 45 control subjects. It seems better to compare the control group with each experimental group separately. Of these seven t-tests the only significant value was that resulting from comparing Group III and the control, where t = 2.14. Again, there seems to be no reason for considering the forgetting in Group III as anything more than a random fluctuation. At the same time it deserves at least passing mention that it is Group III which reached a level of mastery on Task II where response competition between the two tasks might be expected to occur; i.e., mastery equal to or slightly greater than that achieved on Task I. However, there does not seem to be sufficient evidence for proactive inhibition of retention after 24 hours; we conclude that it did not occur.

Returning to Figs. 1 and 2 we can see that all experimental groups continue to be superior to the control. No test for statistical significance is deemed necessary; we conclude that the facilitating effect on the second task of practice on the first task continued after a 24-hour rest and that this is true for all degrees of inter-task similarity. We therefore find proactive facilitation, rather than proactive inhibition, for all conditions.

Finally, let us inspect the differences among the curves for the various experimental groups. Although it appears that the various groups continue to be differentiated by each variable, we may test the differences among the curves for significance. The results of the analysis of variance on total correct responses for the 20 relearning trials are shown in Table 6. It can be seen from the table that both learning and similarity are highly significant; the interaction is not significant. It can be concluded that during relearning following a 24-hour

rest both first-task learning and inter-task similarity continue to differentiate the groups on the second task.

Table 6 - Analysis of Variance on the Total Correct Responses for the 20 Relearning Trials on the Second Task after a 24-Hour Rest

Source	df	Mean Square	F	P
Learning	3	19,471.69	7.49	$<.01$
Similarity	2	13,850.77	5.33	$<.01$
L x S	6	3,498.22	1.35	
Within Groups	174	2,597.18		

Errors: Figure 3 shows mean errors per trial both for the Learning Groups, shown on the left side of the figure, and for the Similarity Groups, shown on the right side. For both sets of groups there are heavy points, to the left of the abcissa break, showing mean errors on the last 10 acquisition trials (trials 51-60 on Task II). The lower curves show retention after 24 hours and include the control curve. The upper curves show retention after 14 months and will be referred to later.

Examination of the 24-hour retention curves in Fig. 3 shows that the Learning Groups did not exhibit increases in errors over the rest and do not differ during relearning. Among the Similarity Groups, Groups B and C show a slight increase in errors over the 24-hour rest. However, there is no indication that further analysis of these error curves would be helpful.

Time: Fig. 4 shows mean time per trial that the left-hand lever was held in the correct position both for the Learning Groups, on the left, and Similarity Groups, on the right. In both sides of the figure the upper sets of curves show time scores during relearning after 24 hours and the lower sets show relearning after 14 months. The figure also includes heavy points showing mean time during the last 10 acquisition trials. The control curve for 24-hour retention is also shown.

Comparison of the 24-hour retention curves for time with those for correct responses (Figs. 1 and 2) shows that the phenomena of forgetting and relearning discussed above with reference to correct re-

sponses occur in much the same way in the time scores. This would be expected since the subject could not record correct responses unless time was being recorded by holding the left lever correctly. Therefore it must be pointed out that the conclusions drawn above from the analysis of correct responses, concerning forgetting, proactive facilitation, etc., apply to the task as a whole and not merely to responses made with the right-hand lever. At the same time, for at least two reasons, one cannot infer that differences in performance measured by correct responses are merely the result of differences in skill with the left lever. In the first place, the correlation between mean time score and mean correct responses for all 186 subjects on the 20 trials of relearning after 24 hours is .652, thus about 42% of the variance in correct responses might be attributed to variance in time score. Secondly, whereas differences among Learning Groups in correct responses might be the result of different amounts of practice with the left lever, this cannot be the interpretation of differences among Similarity Groups, where degree of prior practice is constant. If one were attempting to identify cause and effect, it would be necessary to conclude that differences among Similarity Groups in correct responses were the cause, not the effect, of differences among these groups in time score. Since there seems little point in attempting to attribute causality to either measure, we shall continue to analyze in detail only the correct responses but draw conclusions for the task as a whole.

Retention of the Second Task After 14 Months

Correct Responses: Relearning of the second task after an average of 14 months without practice is shown by the sets of curves in the lower right-hand corner of Figs. 1 and 2. Unfortunately, there is no control group for 14-month retention. By comparing the 14-month curves with the curve for acquisition of the first task, shown in the lower left of Figs. 1 and 2, it can be seen that performance in the 14-month retention test was superior to that of subjects being introduced to the apparatus the first time. It is not known, however, what performance on the second task after 14 months would have been in the absence of practice on the first task.

Forgetting over the long interval was measured by subtracting the score on the first trial of the 14-month test from the mean of the last five trials of the 24-hour retention test (trials 16-20). These data are shown in Table 7. There it may be noted that forgetting tends to increase as degree of learning increases but there is no consistent relationship between similarity and forgetting. The forgetting scores were tested for significance by analysis of variance, the summary table for which is shown in Table 8. The analysis shows that only among the Learning Groups were there significant differences in forgetting. We can conclude that forgetting of the second task over 14 months increased as degree of learning of the first task (and therefore degree of mastery of the second task) increased. However, it may be seen from Table 7 that this effect is due to differences among Groups II, III, and IV; there is no difference between Groups I and II.

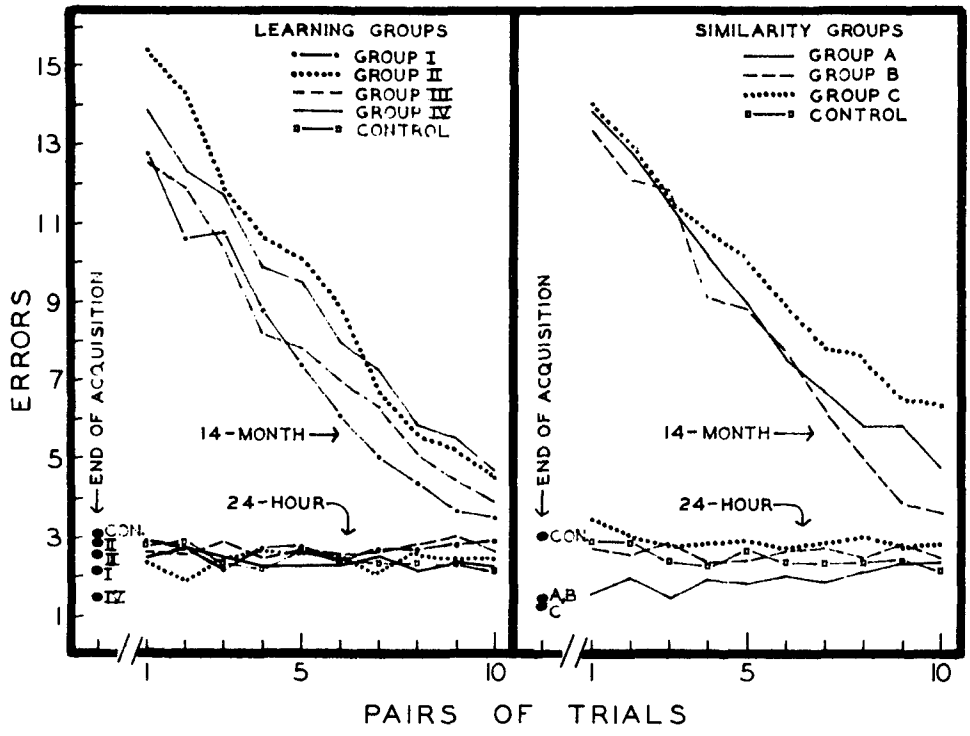

Fig. 3: Errors per trial during relearning of the second task after 24 hours and after 14 months both for the Learning Groups, on the left, and the Similarity Groups on the right.

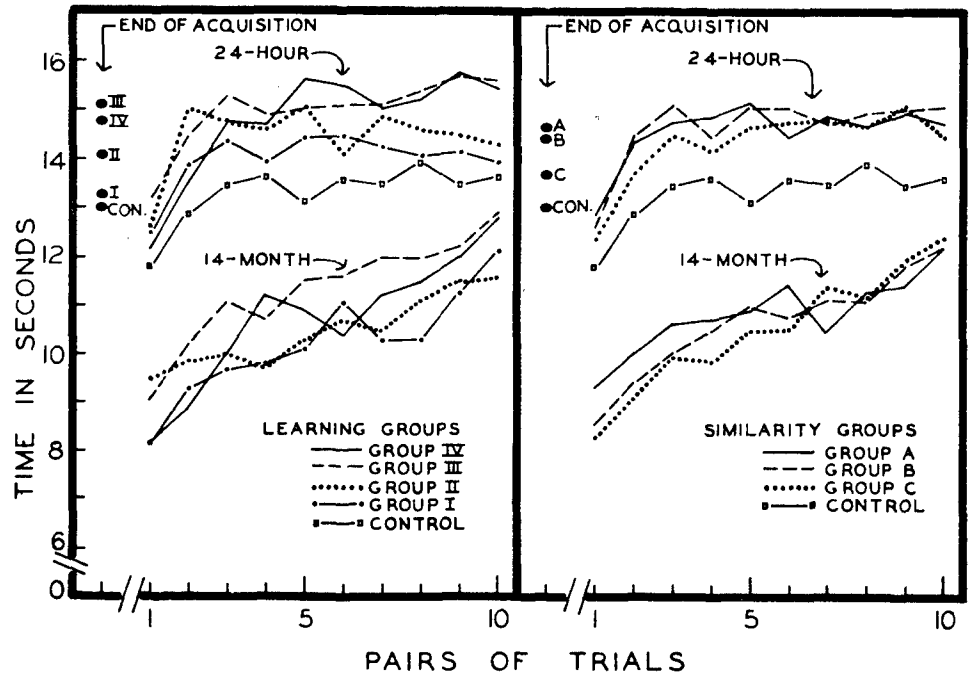

Fig. 4: Time per trial that the left-hand lever was held in correct position during relearning of the second task after 24 hours and after 14 months both for the Learning Groups, on the left, and the Similarity Groups on the right.

Table 7 - Forgetting of the Second Task over 14 Months Measured as the Difference between the Mean of the Last Five Trials Before and the Mean of the First Trial After the 14-Month Interval

Groups	Mean score before rest	Mean score after rest	Difference
I	13.09	2.23	10.86
II	13.84	2.98	10.86
III	15.41	2.80	12.61
IV	15.63	2.58	13.05
A	14.95	2.77	12.18
B	14.79	2.52	12.27
C	13.64	2.64	11.00

Table 8 - Analysis of Variance on Forgetting of the Second Task after 14 Months

Source	df	Mean Square	F	P
Learning	3	61.92	5.50	$<.01$
Similarity	2	32.05	2.85	
L x S	6	5.69		
Within Groups	174	11.25		

In view of the finding that there were significant differences in forgetting over 14 months among the Learning Groups, it is of interest to determine if these groups, as well as the Similarity Groups, actually differed in performance immediately prior to the 14-month interval. Analysis of variance on the last trial (trial 20) of the 24-hour retentinn test resulted in an F of 6.26 for learning and an F of 4.81 for similarity. Both of these are significant at the 1% level. Interaction was not significant. Thus, for the four Learning Groups as a whole there were significant differences in performance immediately prior to the long interval, and significant differences in forgetting over the interval, there being a direct relation between performance and forgetting. This means there has been differential proactive inhibition, resulting from varying amounts of practice on the first task, on retention of the second task after 14 months. It will be recalled that when retention was measured after 24 hours, no proactive inhibition was found. We therefore find support for the contention that proactive inhibition increases as the length of the retention interval increases (6). The Similarity Groups do not yield clear-cut results.

Let us return to Figs. 1 and 2 and examine the course of relearning after 14 months. Relearning is relatively rapid, at least more rapid than original learning of the first task, as can be seen by comparing the slopes of the experimental-group relearning curves with the first-task acquisition curve. However, the experimental groups begin relearning of the second task at a level of mastery lower than they have ever been on this task, as may be seen by noting the level at which they began acquisition of the second task, shown in the left side of the graphs. To state more completely what has occurred: there was relatively large decrement from the end of acquisition of the first task to the beginning of acquisition on the second task (in the first report (3) it was shown that this decrement increased directly both with degree of learning and degree of dissimilarity). In spite of the decrement there was positive transfer to the extent that all groups began acquisition of the second task at a relatively high level. Fourteen months later there is relatively little retention of transfer, if we use recall as a measure; i.e., the level of mastery on the initial relearning trials. Furthermore, there are obviously no differences among either Learning or Similarity Groups at the beginning of relearning. At this point neither variable is differentially retained. However, if we use relearning as the measure, there is, as might be expected, greater retention of transfer, as indicated by the steep slopes of the curves. And finally, differences among the relearning curves suggest that there may be differential retention of transfer. This possibility is tested by the following analyses.

The 14-month retention data were first subjected to the usual double classification analysis of variance, employed several times for previous analyses, in which the score was the total correct responses on all 20 relearning trials. This analysis is summarized in Table 9. The interaction variance closely approaches significance, F = 2.10 with 2.16 needed for significance at the 5% level. This may raise some question as to the appropriate error term for testing the main effects. If the variance within groups is used as error, learning is significant at $P < .05$, as shown in the table, whereas if interaction variance is used

as error, learning is not significant (F = 1.37). Similarity is not significant with either error term.

Table 9 - Analysis of Variance on the Total Correct Responses for All 20 Relearning Trials after 14 Months

Source	df	Mean Square	F	P
Learning	3	5,282.27	2.88	<.05
Similarity	2	664.39		
L x S	6	3,859.96	2.10	
Within Groups	174	1,834.55		

Even though the above analysis might stand by itself, it was felt that a more sensitive test would be obtained by analyzing for trend. For this, Edwards' (4) analysis of variance for trend was modified to handle double classification.[1] The summary of this analysis is shown in Table 10. The appropriate error term for testing learning and similarity and the interaction between the two variables is the variance between subjects in the same group. As may be seen, the F values are the same as those shown in Table 9. However, our main purpose in performing the trend analysis is to test the interactions of the main variables with trials. For these F's the error term is the variance of pooled subjects x trials. Table 10 shows that the interaction of trials and learning is highly significant, but that the interaction of trials and similarity, and the triple interaction, are not significant. The variance for trials is, of course, highly significant. From this analysis we can conclude that degree of learning of the first task is a differentiating variable during relearning of the second task after 14 months, indicated by the significant F's both for learning and for the interaction of trials and learning. Similarity, after 14 months, is no longer a significant variable; all degrees of difference between the two tasks are retained and relearned equally well.

[1] Grateful acknowledgment is made to Miss Jean Paulsen, Department of Psychology, Northwestern University, for developing this method.

Table 10 - Double Classification Analysis of Variance for Trend on Correct Responses for the 20 Relearning Trials after 14 Months

Source	df	Mean Squares	F	P
Learning	3	264.13	2.88	<.05
Similarity	2	33.22		
L x S	6	192.99	2.10	
Trials	19	817.71	140.98	<.01
T x L	57	67.06	11.56	<.01
T x S	38	5.29		
T x L x S	114	5.87	1.01	
Pooled subjects x trials	3306	5.80		
Between subjects in same group	174	91.73		
Total	3719			

Errors: The upper set of curves in Fig. 3 shows errors during relearning. The error curves follow much the same course as the curves for correct responses. In general, Learning Groups showing higher performance levels in terms of correct responses tend to make more errors, whereas among the Similarity Groups better-performing groups tend to make fewer errors. The error data appear to add little to the analysis of results.

Time: The lower set of curves in Fig. 4 shows performance with the left lever during relearning. Again, these time curves are similar to the curves for correct responses. However, the correlation between mean time and mean correct responses over the 20 trials of relearning is only .231, which is significant at the 1% level but not much more. It will be recalled that the correlation during the 24-hour retention test was .652. Why the correlation should drop to .23 after the 14-month interval is not definitely known. The reason, or part of it, may lie in the fact that the variance in correct responses was less during the 14-month test than during the 24-hour test but the corresponding

variances of time scores showed little change. Over the 20 relearning trials after the 24-hour rest, the variance of correct responses was 7.13, of time, 8.18. During the 20 trials of the 14-month test the variance of correct responses was 4.91, of time, 7.88. This drop in the correct-responses variance may or may not be the statistical reason for the drop in correlation, but even if it were, it is not known why the variance should change.

It may also be of interest to compare, insofar as possible, forgetting and relearning of correct responses and time over 14 months. The following method was used to obtain roughly comparable amounts of forgetting. It was first assumed that during the last five trials of the 24-hour retention test both time and correct-response curves were asymptotic, that the means at this point closely represent the upper limit of performance on either measure. For this block of five trials the mean time was 14.92 and the mean correct responses was 14.49. To establish the lower limit of performance, the means were computed for the first five trials of initial exposure to the apparatus, trials 1-5 on the first task. These mean values were 8.42 for time and 2.99 for correct responses. Thus by subtraction the total range of performance is 6.50 for time and 11.50 for correct responses. Then forgetting over the 14-month interval was measured as the difference between the mean of the first five relearning trials on the 14-month test and the asymptotes given above. These forgetting scores are 5.54 for time, 10.22 for correct responses. Thus, a roughly comparable measure of relative forgetting can be obtained by computing the fraction that the forgetting score is of the total range of performance for each response measure. Percentagewise, these values are 85.2% for time and 88.9% for correct responses. Thus, measured as the fraction of the total range of performance, there is only slightly greater forgetting of the right-hand task, i.e., correct responses with the right lever.

The relative amounts of relearning of the two measures may be computed as the percentage that the gain during relearning (from the mean of the first five relearning trials to the last trial) is of the amount forgotten. These values are 61.4% for correct responses, 60.1% for time. Thus the amounts of relearning are very nearly the same.

We may conclude that even though the correlation between correct responses and time scores is low during the 14-month retention test, forgetting and relearning after 14 months differs little for the two measures.

IV. DISCUSSION

The most striking finding of the present study is the consistent effect on acquisition and retention of the final task of different degrees of learning on the first task. Throughout the report we have referred to

many significant differences on the second task as being the result of varying levels of mastery on the first task. It should be clear that we could just as well, perhaps better, have spoken of differences in forgetting and relearning of the second task as being due to differences in level of mastery of the second task. We have preferred to speak of degree of first-task learning because it was only during this task that the amount of practice was varied; the amount of practice was always the same for all subjects throughout acquisition and both relearning sessions on the second task.

Perhaps the most intriguing finding concerns forgetting and relearning over the 14-month period. As indicated by performance on the first trial of relearning after 14 months, there was very great forgetting, a drop to a level of performance never previously experienced on the second task and almost to the level of subjects completely naive to the apparatus. Furthermore, measured at this recall point, neither the learning nor the similarity variable produced differential performance; all groups have nearly identical scores. However, when retention was measured by relearning rather than recall, there was considerable retention; within 20 trials 60% of the amount forgotten was regained. But the interesting point is that, at least among the Learning Groups, "latent" differences emerged during relearning; there was differential retention, as measured by relearning after 14 months, resulting from originally different levels of mastery. And, it should be added, these different levels of mastery resulted, not from varying amounts of practice on the transfer task where retention was being measured, but from differential transfer resulting from varying the degree of learning on the first task. Thus, for this experiment we may speak of differential retention of positive transfer after 14 months without practice.

Perhaps the second major finding of the study is the lack of any difference between various degrees of inter-task similarity after 14 months. It was noted above that in the first report (3) there was some evidence that the effect of similarity was beginning to disappear as early as the latter half of acquisition of the second task. This finding was not, however, as clear-cut as it might have been since, for that original group of 300 subjects, analysis of variance on the 20 relearning trials of their 24-hour retention test resulted in a significant F for similarity. In other words, the original three Similarity Groups of 100 subjects each differed significantly far beyond the 1% level during the first 30 second-task acquisition trials, did not differ during the last 30 acquisition trials, but did differ 24 hours later during the 20 relearning trials. At that time it was pointed out that the significant F for similarity during relearning after 24 hours was due almost entirely to the relatively depressed relearning curve of Group C, the group practicing the least similar task.

The hint one might gather is that with successive periods of practice and rest similarity decreases in effectiveness as a differentiating variable. The results on the 14-month retention test of the present study support this. With the group of 186 subjects (out of

the original 300) used in the present experiment, the three Similarity Groups were significantly different far beyond the 1% level during the first 30 second-task acquisition trials, different just at the 5% level during the last 30 acquisition trials, different at the 1% level during relearning after 24 hours, and not different, actually nearly the same, during relearning after 14 months.

In evaluating the effects of inter-task similarity, as it is defined here, it should be kept in mind that neither stimulus similarity nor response similarity, as they are defined in a paired-associate verbal list (8, 9), was being systematically varied. With the motor task used here the stimuli were identical from task to task. Since the second task was provided by re-pairing stimuli and responses, the responses of the two tasks were highly similar. On the second task the responses were identical in that they were still movements of the same lever in the same plane moving the same distance. Only the direction of movement required for each stimulus was changed. Thus, since we had identical stimuli and highly similar responses for the two tasks, the finding that transfer was positive on all trials, and the finding that positive transfer increased directly with degree of first-task learning, are support for the theory of response generalization (9). This relationship between results with the present apparatus and response-generalization theory was described more fully in the previous report (3), so will not be repeated. Here it may be added that the results with verbal lists apparently indicate that when response similarity is high, the effect of varying first-task (list) learning is greater, measured during second-list acquisition, than when response similarity is low (9). And from the present data, when response similarity is high, varying first-task learning results not only in differential acquisition of the second task but differential retention over long periods of time.

V. IMPLICATIONS FOR TRAINING EQUIPMENT DESIGN

There was one finding in this study which may have considerable significance for the design of training aids and devices. There is currently a great deal of interest in the question of the effectiveness of training equipment as a function of the fidelity with which it simulates the operational situation for which training is intended. The question of degree of simulation required is important since, in general, as fidelity of simulation incorporated in a trainer is increased there is a large, and disproportionate, increase in the cost and time required to produce the device and quite probably an increase in maintenance requirements. It becomes of great practical significance, therefore, to specify minimal requirements for similation consistent with the amount of transfer desired.

This study was not designed to yield quantitative data concerning this problem, but the results obtained with the degree of similarity variable may provide us with a useful hypothesis. It will be remembered that the three Similarity Groups were significantly different during the second-task acquisition trials, were significantly different during re-

learning after 24 hours, but were not significantly different, actually nearly the same, during relearning after 14 months. All of the groups showed positive transfer, as measured by relearning, even after 14 months, but the group having the greatest inter-task similarity showed no more transfer than the group having the least inter-task similarity. This tends to suggest that degree of similarity between training device and operational equipment is of decreasing importance as the interval between training (involving both device and operational equipment) and demand for operational proficiency is increased. Although more research is needed to substantiate and quantify this relationship, we may tentatively recommend that less attention be paid to achieving high fidelity of simulation in situations where a long period of time will intervene between training and the resumption of performance on the operational equipment.

VI. CONCLUSIONS

1. When retention of the positive transfer obtained during acquisition of a transfer task provided on a self-paced discriminative motor device was measured after a 24-hour interval with no practice, there was some forgetting, as measured by recall, but proactive facilitation was found for both recall and relearning.

2. Proactive facilitation after 24 hours increased directly with degree of learning of the training task and with degree of similarity between training and transfer tasks.

3. Immediately preceding a 14-month interval with no practice there were significant differences in performance on the transfer task as a function both of degree of first-task learning and of degree of inter-task similarity.

4. Measured by recall, forgetting of the transfer task over 14 months was so great as to result in performance at the recall point being poorer than at any previous point on the transfer task, not excepting performance on the initial acquisition trial.

5. Measured by recall, forgetting over 14 months increased directly with degree of first-task learning. Thus forgetting increased directly with level of mastery achieved at the point immediately preceding the 14-month interval; there was differential proactive inhibition from the first task on retention of the second task. There was no unambiguous relationship between forgetting and inter-task similarity.

6. Performance at recall after 14 months did not vary as a function of either first-task learning or similarity.

7. Measured by relearning, retention after 14 months was relatively great.

8. Performance during relearning after 14 months varied directly with degree of first-task learning but did not vary with inter-task similarity.

VII. BIBLIOGRAPHY

1. Bunch, M. E. The amount of transfer in rational learning as a function of time. J. comp. Psychol., 1936, 22, 325-337.

2. Bunch, M. E. A comparison of retention and transfer of training from similar material after relatively long intervals of time. J. comp. Psychol., 1941, 32, 217-231.

3. Duncan, C. P. and Underwood, B. J. Transfer of training in motor learning as a function of degree of first-task learning and inter-task similarity. U. S. Air Force Technical Report 52-64, Wright-Patterson Air Force Base, Dayton, Ohio, 1952.

4. Edwards, A. Experimental design in psychological research. New York: Rinehart, 1950.

5. McGeoch, J. A. and Irion, A. L. The psychology of human learning. Second edition. New York: Longman, Green, 1952.

6. Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. J. exp. Psychol., 1948, 38, 29-38.

7. Underwood, B. J. Proactive inhibition as a function of time and degree of prior learning. J. exp. Psychol., 1949, 39, 24-34.

8. Underwood, B. J. Experimental psychology. New York: Appleton-Century-Crofts, 1949.

9. Underwood, B. J. Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. J. exp. Psychol., 1951, 42, 44-53.